On Two-Pair Two-Way Relay Channel with an Intermittently Available Relay

Shih-Chun Lin
Department of ECE, NTUST
Taipei, Taiwan
sclin@mail.ntust.edu.tw

I-Hsiang Wang
Department of EE, NTU
Taipei, Taiwan
ihwang@ntu.edu.tw

***Abstract*—When multiple users share the same resource for physical layer cooperation such as relay terminals in their vicinities, this shared resource may not be always available for every user, and it is critical for transmitting terminals to know whether other users have access to that common resource in order to better utilize it. Failing to learn this critical piece of information may cause severe issues in the design of such cooperative systems. In this paper, we address this problem by investigating a two-pair two-way relay channel with an intermittently available relay. In the model, each pair of users need to exchange their messages within their own pair via the shared relay. The shared relay, however, is only intermittently available for the users to access. The accessing activities of different pairs of users are governed by independent Bernoulli random processes. Our main contribution is the characterization of the capacity region to within a bounded gap in a symmetric setting, for both delayed and instantaneous state information at transmitters. An interesting observation is that the bottleneck for information flow is the quality of state information (delayed or instantaneous) available at the relay, not those at the end users. To the best of our knowledge, our work is the first result regarding how the shared intermittent relay should cooperate with multiple pairs of users in such a two-way cooperative network.**

I. Introduction

Physical layer cooperation has been proposed as a promising approach to increase spectral efficiency, where additional resources are dedicated for cooperation, such as relay terminals in the vicinity. Such resources for cooperation could be shared by many different users. One of the envisioned scenarios for physical layer cooperation is *multi-pair two-way communication via a relay*, where multiple pairs of users exchange their messages within their own pairs, with the help of a relay. The shared resource for cooperation in this scenario is the relay shared by multiple pairs of users. The simplest information theoretic model for studying this problem is the *two-way relay channel* without user-to-user connections. There has been a great deal of works focusing on (multi-pair) two-way relay channels, such as [1], [2]. A conventional assumption in these works is that, the relay is always available for the users to access, so that they can exchange data via the relay all the time.

The work of S.-C. Lin was supported by Ministry of Education, Taiwan under grants "Aiming For the Top University Program" and Ministry of Science and Technology, Taiwan, under Grants MOST 101-2221-E-011-170-MY3. The work of I.-H. Wang was supported by Ministry of Science and Technology, Taiwan, under Grants MOST 103-2221-E-002-089-MY2 and MOST 103-2622-E-002-034.

In practice, however, the opportunity of cooperation may not always exist, mainly because the management and allocation of resources for cooperation (such as relay terminals in their vicinities) lies beyond the physical layer. When multiple users share the same cooperation resource, it may severely impact the design of such cooperative systems if transmitters cannot timely learn how heavily the common resource is currently being utiliized. In the context of multi-pair two-way communication, the issue becomes relevant especially when the spectral activity such as the frequency hopping sequence and/or the frequency coding pattern of a communication link is unknown to a relay which is installed by a third party [3] but shared by multiple pairs of users. Hence, it is of fundamental interest to characterize the capacity of such systems, under various levels of state information availability of other pairs' accessing activities.

In this paper, we take a first step towards this direction by investigating a two-pair two-way relay channel where the two pairs get to access the relay intermittently, under various settings of temporal availability of *activity state information* at transmitters. The availability of accessing the relay is governed by two independent Bernoulli p i.i.d. processes, one for each pair. The terminals can either have delayed information about the activity states, or instantaneous state information. See Figure 1 for an illustration of the channel model.

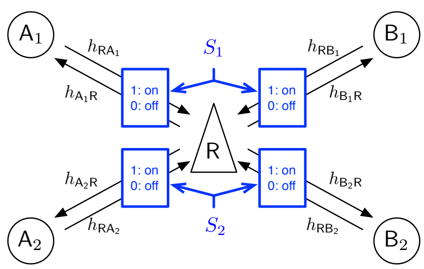


Fig. 1. Two-Pair Two-Way Relay Channel with an Intermittent Relay

Our main contribution is the characterization of the capacity region to within a bounded gap in a symmetric setup, both under the delayed state information setting and the instantaneous state information setting. We show that the two-pair two-

way relay channel can be decomposed into the uplink and the downlink part, and the approximate capacity region is characterized as the intersection of the uplink outer bound region and the downlink outer bound region. The decomposition principle can be viewed as an extension of that in the multi-pair two-way relay channel with a static relay [2]. An interesting observation is hence that the bottleneck for information flow within the system is the quality of state information available at the relay. Towards establishing the achievability of the bound-gap result, for the downlink phase with delayed state information, we have developed a novel scheme that takes care of unequal received signal-to-noise ratios. The scheme complements that in [4] where equal received SNRs are assumed.

We obtain key insights from the binary expansion model [5] for this problem to develop our scheme, where the main novelty is two-fold. First, since the state is not known instantaneously at the relay (transmitter), a lattice-based dirty paper coding (DPC) is employed instead of conventional DPC based on Gaussian random codes. Second, to take care of the unequal received SNRs, instead of quantizing the erased sequences into a single codeword like [4], we propose a successive refinement framework so that stronger receiver can have higher resolution into the quantized signal.

Related work: Two-way relay channel with a static relay has been extensively studied. For the single-pair two-way relay channel, [1] characterized the capacity region to within $\frac{1}{2}$ bit with compute-and-forward [6] and cut-set based outer bound. [2] extended the result to the two-pair two-way relay channel, using insights from the binary-expansion model [5]. However, when the relay is intermittently available, there has been very few results regarding how the shared relay should cooperate with multiple pairs of users. Related works that address intermittence in wireless networks were focused on bursty interference networks. [7] characterized the generalized degrees of freedom of a bursty interference channel with delayed state information and channel output feedback, while [8] [9] studied the degrees of freedom of binary fading interference channels with instantaneous or delayed state information. However, the intermittent availability of cooperation resources have not been investigated widely.

II. Problem Formulation

A. Channel Model

In the system, there are two pairs of *end user* terminals, pair 1: $(\mathsf{A}_1, \mathsf{B}_1)$ and pair 2: $(\mathsf{A}_2, \mathsf{B}_2)$, and one *relay* terminal R. Each terminal can listen and transmit simultaneously, and the blocklength is N. End user U_i in pair i ($\mathsf{U} = \mathsf{A}, \mathsf{B}$, $i = 1, 2$) would like to deliver its message W_{U_i} to the other end user in pair i. The encoding constraints depend on the state information assumption and are detailed in Section II-B.

The two-pair two-way Gaussian relay channel with an intermittent relay is depicted in Figure 1 and defined as follows. The transmitted signals of the five terminals are $X_{\mathsf{A}_1}, X_{\mathsf{B}_1}, X_{\mathsf{A}_2}, X_{\mathsf{B}_2}, X_{\mathsf{R}} \in \mathbb{C}$ respectively, each of which is subject to unit power constraint, and the received signals are

$$Y_{\mathsf{A}_i}[t] = h_{\mathsf{A}_i\mathsf{R}} S_i[t] X_{\mathsf{R}}[t] + Z_{\mathsf{A}_i}[t],\ i = 1, 2,$$
$$Y_{\mathsf{B}_i}[t] = h_{\mathsf{B}_i\mathsf{R}} S_i[t] X_{\mathsf{R}}[t] + Z_{\mathsf{B}_i}[t],\ i = 1, 2, \quad (1)$$
$$Y_{\mathsf{R}}[t] = \sum_{i=1,2} h_{\mathsf{RA}_i} S_i[t] X_{\mathsf{A}_i}[t] + h_{\mathsf{RB}_i} S_i[t] X_{\mathsf{B}_i}[t] + Z_{\mathsf{R}}[t], \quad (2)$$

where the independent additive noises at the five terminals $Z_{\mathsf{A}_1}[t], Z_{\mathsf{B}_1}[t], Z_{\mathsf{A}_2}[t], Z_{\mathsf{B}_2}[t], Z_{\mathsf{R}}[t]$ are $\mathcal{CN}(0, 1)$ i.i.d. over time. $\{S_i[t]\}$ denotes the random process that governs the accessing activity of the two users in pair i, for $i = 1, 2$. $\{S_1[t]\}$ and $\{S_2[t]\}$ are independent Bernoulli p processes, i.i.d. over time [1]. We denote the signal-to-noise ratios as follows: for $i = 1, 2$,

$$\mathsf{SNR}_{\mathsf{RA}_i} := |h_{\mathsf{RA}_i}|^2 \qquad \mathsf{SNR}_{\mathsf{RB}_i} := |h_{\mathsf{RB}_i}|^2$$
$$\mathsf{SNR}_{\mathsf{A}_i\mathsf{R}} := |h_{\mathsf{A}_i\mathsf{R}}|^2 \qquad \mathsf{SNR}_{\mathsf{B}_i\mathsf{R}} := |h_{\mathsf{B}_i\mathsf{R}}|^2$$

Note that we focus the fast fading scenario where a codeword can span over different activity states. This assumption makes our uplink model (2) fundamentally different to the random access channel in [10]. In [10], the slow fading scenario was studied where encoding over different states was prohibited.

B. Activity State Information

We consider two scenarios in this paper regarding how the accessing activity state processes $\{S_1[t]\}$ and $\{S_2[t]\}$ are known to the five terminals, in terms of how the state information helps in encoding.

1) Delayed State Information:

- For end users: for user U_i in pair i ($\mathsf{U} = \mathsf{A}, \mathsf{B}$, $i = 1, 2$), $X_{\mathsf{U}_i}[t] \stackrel{\mathrm{f}}{=} \left(W_{\mathsf{U}_i}, Y_{\mathsf{U}_i}^{t-1}, S_1^{t-1}, S_2^{t-1}\right)$.
- For the relay: $X_{\mathsf{R}}[t] \stackrel{\mathrm{f}}{=} \left(Y_{\mathsf{R}}^{t-1}, S_1^{t-1}, S_2^{t-1}\right)$.

2) Instantaneous State Information:

- For end users: for user U_i in pair i ($\mathsf{U} = \mathsf{A}, \mathsf{B}$, $i = 1, 2$), $X_{\mathsf{U}_i}[t] \stackrel{\mathrm{f}}{=} \left(W_{\mathsf{U}_i}, Y_{\mathsf{U}_i}^{t-1}, S_1^{t}, S_2^{t}\right)$.
- For the relay: $X_{\mathsf{R}}[t] \stackrel{\mathrm{f}}{=} \left(Y_{\mathsf{R}}^{t-1}, S_1^{t}, S_2^{t}\right)$.

The capacity region $\mathscr{C}$ depends on the available activity state information. We take the following notation to denote the capacity region under certain setting of activity state information: $\mathscr{C}(\mathrm{u}, \mathrm{r})$, where the first argument $\mathrm{u} \in \{\mathrm{d}, \mathrm{i}\}$ denotes that the end users have delayed state information (d) or instantaneous state information (i), while the second argument $\mathrm{r} \in \{\mathrm{d}, \mathrm{i}\}$ denotes the type of the available activity state information at the relay terminal.

III. Main Results

In this paper, we focus on the symmetric case where $\mathsf{SNR}_{\mathsf{RU}_i} = \mathsf{SNR}_{\mathsf{R}i}$, $\mathsf{SNR}_{\mathsf{U}_i\mathsf{R}} = \mathsf{SNR}_{i\mathsf{R}}$, for $\mathsf{U} = \mathsf{A}, \mathsf{B}$ and $i = 1, 2$. We focus on characterizing the symmetric rate tuple

[1] In general, the states may be correlated across time and thus allowing us to predict future and improve the throughput. However, discussing the benefit of predicting the future is beyond the scope of this paper, and thus as [7] [8], we impose i.i.d assumptions on states.

(R_1, R_2), where $R_{\mathsf{A}_i} = R_{\mathsf{B}_i} = R_i$ for $i = 1, 2$. Without loss of generality, we assume that $\mathsf{SNR}_{1\mathsf{R}} \geq \mathsf{SNR}_{2\mathsf{R}}$.

To present our main result, let us begin with some definitions useful in characterizing the approximate capacity regions.

Notations:

- Define $\mathcal{C}(x) := \log(1+x)$ (logarithm is of base 2).
- For a $\mathscr{R} \subseteq \mathbb{R}^2$, define the pointwise minus operator $\ominus$ as follows: $\mathscr{R} \ominus (a, b) := \{(x-a, y-b) : (x, y) \in \mathscr{R}\}$.

Uplink Rate Regions: Let $\mathscr{R}^{\text{ul}}_{\text{out}}(\mathrm{d})$ be the collection of $(R_1, R_2) \geq 0$ satisfying

$$\frac{R_1}{p} \leq \mathcal{C}\left(\mathsf{SNR}_{\mathsf{R}1}\right), \quad \frac{R_2}{p} \leq \mathcal{C}\left(\mathsf{SNR}_{\mathsf{R}2}\right), \tag{3}$$

$$\frac{R_1}{p} + \frac{R_2}{p} \leq (1-p)\left(\mathcal{C}\left(\mathsf{SNR}_{\mathsf{R}1}\right) + \mathcal{C}\left(\mathsf{SNR}_{\mathsf{R}2}\right)\right) + p\,\mathcal{C}\left(\mathsf{SNR}_{\mathsf{R}1} + \mathsf{SNR}_{\mathsf{R}2} + 2\sqrt{\mathsf{SNR}_{\mathsf{R}1}\mathsf{SNR}_{\mathsf{R}2}}\right). \tag{4}$$

Let $\mathscr{R}^{\text{ul}}_{\text{in}}(\mathrm{d}) := \mathscr{R}^{\text{ul}}_{\text{out}}(\mathrm{d}) \ominus (1, 1)$. Let $\mathscr{R}^{\text{ul}}_{\text{out}}(\mathrm{i})$ be the collection of $(R_1, R_2) \geq 0$ satisfying (3) – (4) with SNR's replaced by $\frac{\mathsf{SNR}}{p}$, and $\mathscr{R}^{\text{ul}}_{\text{in}}(\mathrm{i})$ be $\mathscr{R}^{\text{ul}}_{\text{in}}(\mathrm{d})$ with SNR's replaced by $\frac{\mathsf{SNR}}{p}$.

Downlink Rate Regions: Let $\mathscr{R}^{\text{dl}}_{\text{out}}(\mathrm{d})$ be the collection of $(R_1, R_2) \geq 0$ satisfying

$$\frac{R_2}{p} \leq \mathcal{C}\left(\mathsf{SNR}_{2\mathsf{R}}\right), \tag{5}$$

$$\frac{R_1}{p} + \frac{R_2}{p(2-p)} \leq \mathcal{C}\left(\mathsf{SNR}_{1\mathsf{R}}\right), \tag{6}$$

$$\frac{R_1}{p(2-p)} + \frac{R_2}{p} \leq \frac{\mathcal{C}(\mathsf{SNR}_{1\mathsf{R}}) - \mathcal{C}(\mathsf{SNR}_{2\mathsf{R}})}{2-p} + \mathcal{C}\left(\mathsf{SNR}_{2\mathsf{R}}\right). \tag{7}$$

Let $\mathscr{R}^{\text{dl}}_{\text{in}}(\mathrm{d}) := \mathscr{R}^{\text{dl}}_{\text{out}}(\mathrm{d}) \ominus (\Delta_1, \Delta_2)$, where

$$\Delta_1 = \frac{p(1-p)}{3-p}\log 3 + \frac{p}{3-p}\log\frac{2\pi e}{12}, \tag{8}$$

$$\Delta_2 = \max\left\{p, \frac{p(1-p)}{3-p}\log 10 + \frac{p}{3-p}\right\}. \tag{9}$$

Let $\mathscr{R}^{\text{dl}}_{\text{out}}(\mathrm{i}) = \mathscr{R}^{\text{dl}}_{\text{in}}(\mathrm{i})$ be the collection of $(R_1, R_2) \geq 0$ with

$$\frac{R_1}{p} \leq \mathcal{C}\left(\frac{\mathsf{SNR}_{1\mathsf{R}}}{p(2-p)}\right), \quad \frac{R_2}{p} \leq \mathcal{C}\left(\frac{\mathsf{SNR}_{2\mathsf{R}}}{p(2-p)}\right),$$

$$\frac{R_1}{p} + \frac{R_2}{p} \leq (1-p)\left(\mathcal{C}\left(\frac{\mathsf{SNR}_{1\mathsf{R}}}{p(2-p)}\right) + \mathcal{C}\left(\frac{\mathsf{SNR}_{2\mathsf{R}}}{p(2-p)}\right)\right) + p\,\mathcal{C}\left(\frac{\mathsf{SNR}_{1\mathsf{R}} + \mathsf{SNR}_{2\mathsf{R}}}{p(2-p)}\right).$$

Before we proceed, we provide some numerical evaluations to illustrate various regions defined above. We set the on/off probability of activity state $p = 0.6$. In Figure 2(a), we show $\mathscr{R}^{\text{ul}}_{\text{in}}(\mathrm{d})$ and $\mathscr{R}^{\text{ul}}_{\text{in}}(\mathrm{i})$ with $\mathsf{SNR}_{\mathsf{R}1} = 30 + 20\log 1.5$ dB and $\mathsf{SNR}_{\mathsf{R}2} = 30$ dB. In Figure 2(b), we show $\mathscr{R}^{\text{dl}}_{\text{in}}(\mathrm{d})$ and $\mathscr{R}^{\text{dl}}_{\text{in}}(\mathrm{i})$ with $\mathsf{SNR}_{1\mathsf{R}} = 30 + 20\log 1.5$ dB and $\mathsf{SNR}_{2\mathsf{R}} = 30$ dB.

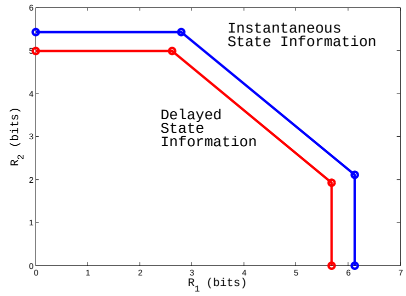
Remark: Note that $\mathscr{R}^{\text{dl}}_{\text{in}}(\mathrm{d})$ is the smallest among four regions in Figure 2(a) and 2(b). If the relay has only delayed state information, the downlink from relay becomes the bottleneck for information flow within the system.

Our main result is summarized in the following theorem.

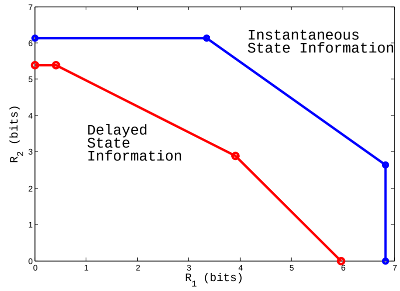
Theorem 3.1 (Capacity Region to within a Bounded Gap): For capacity region $\mathscr{C}(\mathrm{u}, \mathrm{r})$, we have inner and outer bounds

$$\mathscr{C}(\mathrm{u}, \mathrm{r}) \supseteq \mathscr{R}^{\text{ul}}_{\text{in}}(\mathrm{u}) \cap \mathscr{R}^{\text{dl}}_{\text{in}}(\mathrm{r}), \ \forall\, (\mathrm{u}, \mathrm{r}) \in \{\mathrm{d}, \mathrm{i}\}^2, \tag{10}$$

$$\mathscr{C}(\mathrm{u}, \mathrm{r}) \subseteq \mathscr{R}^{\text{ul}}_{\text{out}}(\mathrm{u}) \cap \mathscr{R}^{\text{dl}}_{\text{out}}(\mathrm{r}), \forall\, (\mathrm{u}, \mathrm{r}) \in \{\mathrm{d}, \mathrm{i}\}^2. \tag{11}$$



(a)



(b)

Fig. 2. Bounded-gap uplink (a) and downlink (b) inner bound regions with delayed and instantaneous activity state information.

Since for all $(\mathrm{u}, \mathrm{r}) \in \{\mathrm{d}, \mathrm{i}\}^2$, $\mathscr{R}^{\text{ul}}_{\text{out}}(\mathrm{u})$ and $\mathscr{R}^{\text{ul}}_{\text{in}}(\mathrm{u})$ are within a bounded gap, and so are $\mathscr{R}^{\text{dl}}_{\text{out}}(\mathrm{r})$ and $\mathscr{R}^{\text{dl}}_{\text{in}}(\mathrm{r})$, we have characterized the capacity region to within a bounded gap.

Proof: Regarding the proof of the converse, we employ cut-set based outer bounds and enhance the downlink channel to a degraded broadcast channel where feedback does not increase the capacity region [11], [12]. Details can be found in the Appendix.

Regarding the achievability, here we provide the scheme for the inner bound of $\mathscr{C}(\mathrm{d}, \mathrm{d})$ in (10), the case where end users and relay all have delayed state information. The proofs for the other three combinations in Theorem 3.1 easily follow, and are also provided in the Appendix.

Our scheme consists of two phases: the uplink phase and the downlink phase. In the uplink phase, the relay terminal aims to decode the two XORs of the two pairs of messages $\Sigma_i = W_{\mathsf{A}_i} \oplus W_{\mathsf{B}_i}$, $i = 1, 2$ from its received signal, and store them for later uses. Hence, it can be viewed as a function computation problem over a multiple access channel. In the downlink phase, the relay terminal re-encodes the stored XORs $\{\Sigma_1, \Sigma_2\}$ and delivers Σ_i to end users $\{\mathsf{A}_i, \mathsf{B}_i\}$ for $i = 1, 2$. The end user terminals decode their desired messages from the XORs by using its self message as side information. Hence, it can be viewed as a broadcast channel with two independent messages $\{\Sigma_1, \Sigma_2\}$ and four receivers $\{\mathsf{A}_1, \mathsf{B}_1, \mathsf{A}_2, \mathsf{B}_2\}$, where $\{\mathsf{A}_i, \mathsf{B}_i\}$ aim to decode Σ_i, for $i = 1, 2$.

Further note that in the symmetric setting, since the rate of the messages W_{A_i} and W_{B_i} are both R_i, the rate of the XOR Σ_i is also R_i, for $i = 1, 2$. Hence, we are able to establish the inner bound region of achievable (R_1, R_2) as the intersection

of the inner bound region of the uplink phase and that of the downlink phase, denoted by $\mathscr{R}_{\text{in}}^{\text{ul}}$ and $\mathscr{R}_{\text{in}}^{\text{dl}}$ respectively. Below we give the proof sketches for the uplink phase in Sec. IV and the downlink phase in Sec. V. The detailed proofs are given in the Appendix. ■

IV. Proof Sketch of the Inner Bound $\mathscr{R}_{\text{in}}^{\text{ul}}(\text{d})$ in (10) for Uplink with Delayed State Information

To achieve $\mathscr{R}_{\text{in}}^{\text{ul}}(\text{d})$ in the uplink phase, we will use lattice-based compute-and-forward [6]. Casting it as a function computation problem over the four-transmitter multiple access channel, the relay can successfully decode the XORs of messages $\Sigma_i = W_{\mathsf{A}_i} \oplus W_{\mathsf{B}_i}$, $i = 1, 2$ from its received signal (2) without explicitly decoding the four messages $\{W_{\mathsf{A}_1}, W_{\mathsf{B}_1}, W_{\mathsf{A}_2}, W_{\mathsf{B}_2}\}$, thanks to the linearity of lattice codes. Compared with the scheme in [2], our scheme needs to deal with the additional ergodic activity states $\{S_1(t), S_2(t)\}$ and the delayed state information. Also we adopt joint lattice deocding from [13], which has better performance than the successive lattice decoding in [2].

The details of achieving $\mathscr{R}_{\text{in}}^{\text{ul}}(\text{d})$ in (10) come as follows. First, we assume that the channel gains in (2) are real, which is without loss of generality since we can pre-rotate the phase of the complex channel before transmission. Then we collect the real and imaginary parts of the T received symbols at the relay as [13], and focus on the following real equivalent uplink channel from (2) as

$$\mathbf{y}_{\mathsf{R}} = [\mathbf{H}_{A_1}\ \mathbf{H}_{A_2}] \begin{bmatrix} \mathbf{x}_{\mathsf{A}_1} + \mathbf{x}_{\mathsf{B}_1} \\ \mathbf{x}_{\mathsf{A}_2} + \mathbf{x}_{\mathsf{B}_2} \end{bmatrix} + \mathbf{z}_{\mathsf{R}}, \tag{12}$$

where the $2T \times 1$ real vector $\mathbf{y}_{\mathsf{R}}$ is formed from $Y_{\mathsf{R}}[t]$ as

$$\mathbf{y}_{\mathsf{R}} = [\text{Re}(Y_{\mathsf{R}}[1]), \text{Im}(Y_{\mathsf{R}}[1]), \ldots, \text{Re}(Y_{\mathsf{R}}[T]), \text{Im}(Y_{\mathsf{R}}[T])]^T,$$

and $2T \times 1$ $\mathbf{x}_{\mathsf{A}_1}, \mathbf{x}_{\mathsf{B}_1}, \mathbf{x}_{\mathsf{A}_2}, \mathbf{x}_{\mathsf{B}_2}, \mathbf{z}_{\mathsf{R}}$ are similarly formed from $X_{\mathsf{A}_1}[t], X_{\mathsf{B}_1}[t], X_{\mathsf{A}_2}[t], X_{\mathsf{B}_2}[t], Z_{\mathsf{R}}[t]$ respectively. The $2T \times 2T$ diagonal channel matrix for pair i is

$$\mathbf{H}_{\mathsf{A}_i} = |h_{\mathsf{RA}_i}| \cdot \text{diag}(S_i(1), S_i(1), \ldots, S_i(T), S_i(T)). \tag{13}$$

The transmitted vector for user U_i in pair i ($\mathsf{U} = \mathsf{A}, \mathsf{B}$, $i = 1, 2$) is

$$\mathbf{x}_{\mathsf{U}_i} = ([\mathbf{c}_{\mathsf{U}_i} - \mathbf{d}_{\mathsf{U}_i}]) \mod \Lambda_S. \tag{14}$$

With Λ_i being the coding lattice [13] [14], the message W_{U_i} is encoded using lattice codeword $\mathbf{c}_{\mathsf{U}_i} \in \Lambda_i$, and the shaping lattice $\Lambda_S \subset \Lambda_i$. As [13] [14], the independent dither $\mathbf{d}_{\mathsf{U}_i}$ is uniformly distributed in the Voronoi region of the shaping lattice Λ_S, and $\mod \Lambda_S$ is the modulo-lattice operation. At the relay, it performs joint lattice decoding for XORs Σ_1 and Σ_2 on the following post-processed received signal

$$\left(\mathbf{W}\mathbf{y}_{\mathsf{R}} + [(\mathbf{d}_{\mathsf{A}_1} + \mathbf{d}_{\mathsf{B}_1})^T\ \ (\mathbf{d}_{\mathsf{A}_2} + \mathbf{d}_{\mathsf{B}_2})^T]^T\right) \mod (\Lambda_S \times \Lambda_S).$$

From (12), by choosing $\mathbf{W} = 2\mathbf{H}^T(2\mathbf{H}\mathbf{H}^T + \mathbf{I})^{-1}$ where $\mathbf{H} = [\mathbf{H}_{\mathsf{A}_1}\ \mathbf{H}_{\mathsf{A}_2}]$, the achievable sum rate $R_1/p + R_2/p$ has gap $2/p$ to the RHS of (4). The other two rate constraints for $\mathscr{R}_{\text{in}}^{\text{ul}}(\text{d})$ can be similarly proved to be achievable.

V. Proof Sketch of the Inner Bound $\mathscr{R}_{\text{in}}^{\text{dl}}(\text{d})$ in (10) for Downlink with Delayed State Information

In our symmetric setting, since A_i and B_i have the same receiver SNRs and are under the same activity state $\{S_i[t]\}$, for $i = 1, 2$, we can treat the downlink as a broadcast channel (1) where the relay sends Σ_1 to user B_1 and Σ_2 to user B_2 respectively, with delayed state information. Compared with [4], which is focused on ergodic Rayleigh fading downlink with equal received SNRs, our downlink (1) has different on/off channel statistics and non-equal $\mathsf{SNR}_{1\mathsf{R}} \geq \mathsf{SNR}_{2\mathsf{R}}$. These two differences raise new challenges for obtaining bounded-gap capacity results.

For the corner point of the outer bound region where (7) and (5) intersect, achieving it to within a bounded gap can be simply done by Gaussian superposition coding. Thus we focus on the other corner point where (6) and (7) intersect:

$$R_1 = p\left(\mathcal{C}(\mathsf{SNR}_{1\mathsf{R}}) - \mathcal{C}(\mathsf{SNR}_{2\mathsf{R}})\right) + \tfrac{p(2-p)}{3-p}\mathcal{C}(\mathsf{SNR}_{2\mathsf{R}}), \tag{15}$$

$$R_2 = \tfrac{p(2-p)}{3-p}\mathcal{C}(\mathsf{SNR}_{2\mathsf{R}}). \tag{16}$$

Our scheme to achieve $(R_1 - \Delta_1, R_2 - \Delta_2)$ with R_1, R_2 taken from (15)(16) is a non-trivial extension of the scheme in the binary erasure broadcast channel [11], where (Δ_1, Δ_2) are given in (8)(9). To obtain insights, we start with a binary-expansion model [5] for this problem as follows.

A. Insights from Binary-Expansion Model

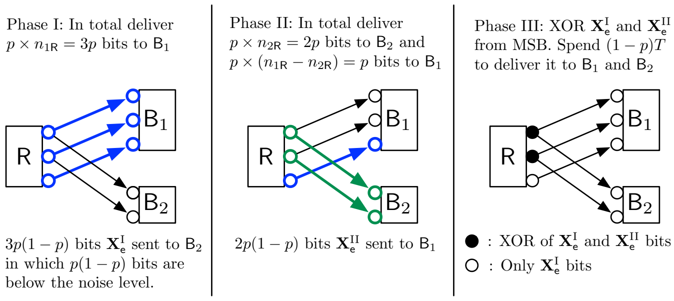


Fig. 3. Example for achieving corner point (17) with $(n_{1\mathsf{R}}, n_{2\mathsf{R}}) = (3, 2)$ for the binary expansion downlink with delayed state information

In this subsection, we employ a binary expansion model corresponding to the downlink phase (1) to obtain insights. In this model, the transmitted and received signals are binary vectors in $\mathbb{F}_2^q$, where $\mathbb{F}_2$ denotes the binary field $\{0, 1\}$. The received signals are $Y_{\mathsf{B}_i}[t] = \mathbf{H}_{\mathsf{B}_i\mathsf{R}} S_i[t] X_{\mathsf{R}}[t]$, $i = 1, 2$, where additions are modulo-two component-wise. Channel transfer matrices are defined as follows: for $i = 1, 2$, $\mathbf{H}_{\mathsf{B}_i\mathsf{R}} := \mathbf{S}^{q - n_{i\mathsf{R}}}$ where $q = \max_{i=1,2}\{n_{i\mathsf{R}}\}$ and $\mathbf{S} \in \mathbb{F}_2^{q \times q}$ is the shift matrix defined in [5]. The corner point corresponds to (15) and (16) in this model is

$$(R_1, R_2) = \left(p(n_{1\mathsf{R}} - n_{2\mathsf{R}}) + \tfrac{p(2-p)}{3-p}n_{2\mathsf{R}}, \tfrac{p(2-p)}{3-p}n_{2\mathsf{R}}\right). \tag{17}$$

To achieve this point, the relay uses a three-phase coding scheme extending that in [11]. In Phases I and II (each with block length T), the relay sends bits intended for B_1 and B_2, using the top $n_{1\mathsf{R}}$ and $n_{2\mathsf{R}}$ levels respectively. In addition, in

Phase II the relay also uses the bottom $(n_{1\mathsf{R}} - n_{2\mathsf{R}})$ levels to deliver additional bits to B_1. Hence, B_1 and B_2 receive roughly $Tp\,(n_{1\mathsf{R}} + n_{1\mathsf{R}} - n_{2\mathsf{R}})$ and $Tpn_{2\mathsf{R}}$ desired bits in Phase I and II respectively.

In Phase I, there will be roughly $Tp(1-p)n_{1\mathsf{R}}$ bits which are erased at B_1 but erroneously sent to B_2 can be used as side-information. We denote this length-$Tp(1-p)$ sequence of $n_{1\mathsf{R}}$-level binary vector by $\mathbf{X}_{\mathsf{e}}^{\mathrm{I}}$. Note that the bottom $(n_{1\mathsf{R}} - n_{2\mathsf{R}})$ levels will lie *below* the noise level at B_2 and will NOT appear in this binary expansion model. Similarly in Phase II, there will be such a length $Tp(1-p)$ sequence of $n_{2\mathsf{R}}$-level binary vector intended for B_2 but only received by B_1. We denote it by $\mathbf{X}_{\mathsf{e}}^{\mathrm{II}}$. We aim to *recycle* these bits in Phase III.

The block length of Phase III is roughly $Tp(1-p)$. In Phase III, the relay makes use of delayed state information to form $\mathbf{X}_{\mathsf{e}}^{\mathrm{I}}$ and $\mathbf{X}_{\mathsf{e}}^{\mathrm{II}}$. Then it sends out $\mathbf{X}_{\mathsf{e}}^{\mathrm{I}} \oplus \mathbf{X}_{\mathsf{e}}^{\mathrm{II}}$ from the MSB level as depicted on the rightmost of Figure 3. Hence the bottom $(n_{1\mathsf{R}} - n_{2\mathsf{R}})$ levels consists of bits in the bottom levels of $\mathbf{X}_{\mathsf{e}}^{\mathrm{I}}$ only. With side information received in Phase I and II, each receiver can decode the desired bits from the received XORs. In total the numbers of bits *recycled* in this phase are $Tp(1-p)n_{1\mathsf{R}}$ and $Tp(1-p)n_{2\mathsf{R}}$ at B_1 and B_2 respectively.

Putting everything together, we achieve $R_2 = \frac{p(2-p)}{3-p} n_{2\mathsf{R}}$ and $R_1 = \frac{p\{(3-p)n_{1\mathsf{R}} - n_{2\mathsf{R}}\}}{3-p} = p(n_{1\mathsf{R}} - n_{2\mathsf{R}}) + \frac{p(2-p)}{3-p} n_{2\mathsf{R}}$.

B. Proof Sketch of Bounded-gap Achievement to the Corner Point (15) (16) *of the Outer Bound Region*

Extending to the Gaussian case, we face the following two challenges. First, in Gaussian channel, we are sending complex symbols instead of binary bits and there will be additive Gaussian noise. Second, we need to incorporate superposition coding into Phase II of Fig. 3, while B_1 may not be able to decode and cancel the higher-layer codeword since the erasure state process at B_1 and B_2 are different. Note that only signals of B_2 in Phase II have recycling from Phase III.

We solve the first challenge by resending erased symbols instead of bits in the third phase. To do this, the relay will *quantize* the sum sequence formed by the erased symbols, $X_{\mathsf{e}}^{\mathrm{I}} + X_{\mathsf{e}}^{\mathrm{II}}$, and then send out the quantization indices. Based on the insight learned in the binary expansion model, we know that the resolution of reconstruction must be different: B_1 requires higher resolution than B_2 since $X_{\mathsf{e}}^{\mathrm{I}}$ goes deeper in the bit levels. Hence, instead of directly quantizing into a single quantization index, we employ *successive refinement* source coding [15] so that B_1 is able to get a higher resolution in reconstruction. Again gaining insights from the binary expansion model, since the number of layers used by B_i is $n_{i\mathsf{R}}$ for $i = 1, 2$ in Phase I and II respectively, the MSE of the reconstruction at B_i should be inverse proportional to $\mathsf{SNR}_{i\mathsf{R}}, i = 1, 2$.

For the second challenge, we aim to solve it using dirty paper coding (DPC). However, the conventional DPC requires fully known channel information $S_1(t)h_{\mathsf{B}_1\mathsf{R}}$ at the transmitter [14]. In our case, the current on/off state $S_1(t)$ is unknown at the relay. Hence, we propose new one-dimensional (symbol-based) lattice strategy to solve this problem.

Our scheme is summarized as follows

Phase I: By using random Gaussian codebook, relay sends coded symbols $X_{\mathsf{R}}[t]$, $t = 1 \ldots T$ from the codeword representing message for user B_1.

Phase II: Relay sends $X_{\mathsf{R}}[t] = X_{2\mathsf{R}}[t] + X_{1\mathsf{R}}[t]$, $t = T + 1 \ldots 2T$, where $X_{2\mathsf{R}}[t]$ are coded symbols for user B_2 and

$$X_{1\mathsf{R}}[t] = (C_{1\mathsf{R}}[t] - w \cdot h_{\mathsf{B}_1\mathsf{R}} X_{2\mathsf{R}}[t] - d[t]) \mod L \quad (18)$$

where similar to (14), $C_{1\mathsf{R}}[t]$ is coded symbol for user B_1. $d[t]$ is the independent dither. For a real number $x, x \mod L = x - Q_L(x)$ with $Q_L(x)$ being the nearest multiple of L to x.

Phase III: Let the erased symbols sent to the wrong receiver in Phase 1 and 2 be $X_{\mathsf{e}}^{\mathrm{I}}$ and $X_{\mathsf{e}}^{\mathrm{II}}$ respectively. Relay first quantizes the length $Tp(1-p)$ sequence $X_{\mathsf{e}}^{\mathrm{I}} + X_{\mathsf{e}}^{\mathrm{II}}$ using successive refinement into indexes i_{c} and i_{r}, where i_{c} is the common index which will be decoded for both B_1 and B_2 while i_{r} is the refinement index which will be decoded only at B_1. Gaussian superposition channel coding with length $T(1-p)$ is adopted to transmit $(i_{\mathsf{c}}, i_{\mathsf{r}})$.

Now, user B_2 can know the noisy reconstruction $X_{\mathsf{e}}^{\mathrm{I}} + X_{\mathsf{e}}^{\mathrm{II}} + Z_{\mathsf{D}2}$ with MSE D_2, by decoding i_{c}. With proper power allocation, B_1 (better channel) knows the reconstruction $X_{\mathsf{e}}^{\mathrm{I}} + X_{\mathsf{e}}^{\mathrm{II}} + Z_{\mathsf{D}1}$ by successively decoding i_{c} and i_{r}, where reconstruction error $Z_{\mathsf{D}1}$ has smaller MSE D_1 than that of $Z_{\mathsf{D}2}$. For this two-receiver source-channel coding, the rates for common index i_{c} and refinement index i_{r} are chosen as

$$\log\left(1 + \tfrac{2}{\mathsf{D}_2}\right) \text{ and } \log\left(1 + \tfrac{2}{\mathsf{D}_1}\right) - \log\left(1 + \tfrac{2}{\mathsf{D}_2}\right) \quad (19)$$

respectively. To ensure successful channel decoding at receivers, we need to carefully choosing the power allocation of the superposition channel coding, as well as D_1 and D_2 in (19). Let the power allocation for indexes i_{c} and i_{r} be $\mathsf{SNR}_{\mathsf{c}}$ and $\mathsf{SNR}_{\mathsf{r}}$ respectively. We choose $\mathsf{SNR}_{\mathsf{r}} = 1/\mathsf{SNR}_{2\mathsf{R}}, \mathsf{SNR}_{\mathsf{c}} = 1 - \mathsf{SNR}_{\mathsf{r}}$. (Here we only provide the proof when $\mathsf{SNR}_{2\mathsf{R}} \geq 2$, since the bounded-gap result for $\mathsf{SNR}_{2\mathsf{R}} < 2$ is trivial.) For B_2 to correctly decode i_{c}, from (19) and the lengths of channel and source codes, we need to choose

$$\mathsf{D}_2 = \tfrac{4}{\mathsf{SNR}_{2\mathsf{R}} - 1} \quad (20)$$

For receiver B_1 to decode both i_{c} and i_{r}, we choose

$$\mathsf{D}_1 = \tfrac{4}{\mathsf{SNR}_{1\mathsf{R}} + \mathsf{SNR}_{2\mathsf{R}}}. \quad (21)$$

Note that D_i is inverse proportional to $\mathsf{SNR}_{i\mathsf{R}}, i = 1, 2$, consistent with the insights from the binary expansion model.

Now receivers $\mathsf{B}_i, i = 1, 2$ can obtain reconstructions of erased symbols $X_{\mathsf{e}}^{\mathrm{I}} + X_{\mathsf{e}}^{\mathrm{II}}$ with D_1 in (21) and D_2 in (20) respectively. With side-information $X_{\mathsf{e}}^{\mathrm{I}}$ (noisy) from Phase I, B_2 can combine $Tp(1-p)$ reconstructed symbols $X_{\mathsf{e}}^{\mathrm{II}} + Z_{\mathsf{D}2} - Z_{\mathsf{B}_2}$ and the Tp un-erased symbols received in Phase II to decode XOR Σ_2. Then (16) is achievable with bounded gap Δ_2. To see this, we can first upper-bound the MSE distortion D_2 in (20) as

$$\mathsf{D}_2 \leq \tfrac{8}{\mathsf{SNR}_{2\mathsf{R}}}. \quad (22)$$

By choosing the power allocation of $X_{1\mathsf{R}}$ and $X_{2\mathsf{R}}$ in Phase II be $1/\mathsf{SNR}_{2\mathsf{R}}$ and $1 - 1/\mathsf{SNR}_{2\mathsf{R}}$ respectively, together with

the independence of these two signals, we have the following achievable rate for user B_2

$$R_2 \geq \frac{1}{(3-p)}\left(p(1-p)\mathcal{C}\left(\frac{1-\frac{1}{\mathsf{SNR}_{2\mathsf{R}}}}{\frac{8}{\mathsf{SNR}_{2\mathsf{R}}}+\frac{1}{\mathsf{SNR}_{2\mathsf{R}}}+\frac{1}{\mathsf{SNR}_{2\mathsf{R}}}}\right)\right.$$
$$\left.+\,p\,\mathcal{C}\left(\frac{1-\frac{1}{\mathsf{SNR}_{2\mathsf{R}}}}{\frac{1}{\mathsf{SNR}_{2\mathsf{R}}}+\frac{1}{\mathsf{SNR}_{2\mathsf{R}}}}\right)\right) \quad (23)$$

where (22) is applied to obtain the first term in the RHS of (23). Then bounded gap result can be obtained from (23).

Now we show that for user B_1, rate $R_1-\Delta_1$, with R_1 in (15), is achievable. Following similar procedure as B_2 aforementioned, by combining the erased symbols with the un-erased symbols received in Phase I, the following rate is achievable to decode XOR Σ_1,

$$\tfrac{p(2-p)}{3-p}\mathcal{C}(\mathsf{SNR}_{1\mathsf{R}}) - \tfrac{p(1-p)}{3-p}\log(3). \quad (24)$$

where the following inequality from (21) is used

$$\mathsf{D}_1 \leq \tfrac{2}{\mathsf{SNR}_{1\mathsf{R}}}.$$

Moreover, user B1 can decode additional messages by forming the following channel from the un-erased symbols in Phase II,

$$[C_{1\mathsf{R}}[t]+E_L(t)] \mod L, \quad (25)$$

where $E_L(t)=(wh_{\mathsf{B}_1\mathsf{R}}-1)X_{1\mathsf{R}}[t]+wZ_{\mathsf{B}_1}[t]$. The channel (25) is a modulo-L channel with length Tp and power $L^2/12 = 1/\mathsf{SNR}_{2\mathsf{R}}$, then rate

$$\tfrac{p}{3-p}\left(\log\left(\tfrac{\mathsf{SNR}_{1\mathsf{R}}}{\mathsf{SNR}_{2\mathsf{R}}}\right)-\log(2\pi e/12)\right) \quad (26)$$

is achievable. By summing (26) and (24), our achievable rate for user B_1 has bounded gap to (15)

Appendix

A. *Detailed proof of the Inner Bound $\mathscr{R}^{\mathrm{ul}}_{\mathrm{G,in}}(\mathrm{d})$ in (10)*

Here we focus on the bounded-gap achievability to the sum rate in (4) by joint lattice decoding. From (12) and (14), it can be easily shown that the post-processed signal

$$\left(\mathbf{W}\mathbf{y}_{\mathsf{R}}+[(\mathbf{d}_{\mathsf{A}_1}+\mathbf{d}_{\mathsf{B}_1})^T\ \ (\mathbf{d}_{\mathsf{A}_2}+\mathbf{d}_{\mathsf{B}_2})^T]^T\right) \bmod (\Lambda_S\times\Lambda_S).$$

equals to

$$\left(\left[(\mathbf{c}_{A_1}+\mathbf{c}_{B_1})^T\ \ (\mathbf{c}_{A_2}+\mathbf{c}_{B_2})^T\right]^T+\mathbf{E}\right) \bmod (\Lambda_S\times\Lambda_S), \quad (27)$$

where

$$\mathbf{E}=(\mathbf{WH}-\mathbf{I})\left[(\mathbf{x}_{\mathsf{A}_1}+\mathbf{x}_{\mathsf{B}_1})^T\ \ (\mathbf{x}_{\mathsf{A}_2}+\mathbf{x}_{\mathsf{B}_2})^T\right]^T+\mathbf{W}\mathbf{z}_{\mathsf{R}}, \quad (28)$$

and $\mathbf{H}=[\mathbf{H}_{\mathsf{A}_1}\ \mathbf{H}_{\mathsf{A}_2}]$. Moreover, in (27), the sum of user codewords within a pair $i=1,2$ is still a lattice codeword

$$(\mathbf{c}_{\mathsf{A}i}+\mathbf{c}_{\mathsf{B}i}) \bmod \Lambda_S \in \Lambda_i.$$

Then from [13], the following sum rate is achievable by jointly lattice decoding

$$R_1+R_2 \geq \lim_{T\to\infty}\frac{1}{2T}\log\left(\frac{|\frac{1}{2}\mathbf{I}|}{|\Sigma_E|}\right), \quad (29)$$

where Σ_E is the covariance matrix of $\mathbf{E}$ in (28). With $\mathbf{W}$ chosen as the MMSE filter, the information lossless property of MMSE estimation can be invoked, then the sum rate in (29) becomes

$$R_1+R_2 \geq \lim_{T\to\infty}\frac{1}{2T}\log\left(\left|\frac{1}{2}\mathbf{I}+\mathbf{H}\mathbf{H}^T\right|\right). \quad (30)$$

Now from (13) ($\mathbf{H}=[\mathbf{H}_{\mathsf{A}_1}\ \mathbf{H}_{\mathsf{A}_2}]$) and the ergodicity of state process,

$$\begin{aligned}R_1+R_2 \geq &\mathrm{E}_{S_1,S_2}\left[\mathcal{C}\left(S_1\mathsf{SNR}_{\mathsf{R}1}+S_2\mathsf{SNR}_{\mathsf{R}2}\right)\right]-1\\ &\geq p(1-p)\left(\mathcal{C}\left(\mathsf{SNR}_{\mathsf{R}1}+\mathcal{C}\mathsf{SNR}_{\mathsf{R}2}\right)\right)\\ &\quad+p^2\mathcal{C}\left(\mathsf{SNR}_{\mathsf{R}1}+\mathsf{SNR}_{\mathsf{R}2}\right)-1 \quad (31)\end{aligned}$$

To compared (31) with (4), it can be easily checked that

$$\begin{aligned}&\mathcal{C}\left(\mathsf{SNR}_{\mathsf{R}1}+\mathsf{SNR}_{\mathsf{R}2}+2\sqrt{\mathsf{SNR}_{\mathsf{R}1}\mathsf{SNR}_{\mathsf{R}2}}\right)\\ &\leq 1+\mathcal{C}\left(\mathsf{SNR}_{\mathsf{R}1}+\mathsf{SNR}_{\mathsf{R}2}\right). \quad (32)\end{aligned}$$

Then the bounded-gap result for sum rate (4) is established. The bounded-gap achievement to the RHSs of (3) follows similarly. As a final note, decoding correct $(\mathbf{c}_{\mathsf{A}i}+\mathbf{c}_{\mathsf{B}i}) \bmod \Lambda_S$ from aforementioned joint lattice decoding equals to decode corrects XORs $\Sigma_i=W_{\mathsf{A}_i}\oplus W_{\mathsf{B}_i}, i=1,2$ [6].

B. *Detailed proof of the inner bound $\mathscr{R}^{\mathrm{dl}}_{\mathrm{G,in}}(\mathrm{d})$ in (10)*

Here we provide the detailed proof for achieving the Gaussian downlink rate region with delayed state information $\mathscr{R}^{\mathrm{dl}}_{\mathrm{G,in}}(\mathrm{d})$. We assume that state (S_1^{t-1},S_2^{t-1}) are known at both receivers B_1 and B_2 at time t. In the following, we only provide the proof when $\mathsf{SNR}_{2\mathsf{R}}\geq 2$, since the bounded-gap result for $\mathsf{SNR}_{2\mathsf{R}}<2$ is trivial. We first focus on the proposed three-phase scheme to achieve $(R_1-\Delta_1, R_2-\Delta_2)$ from (15)(16), where (Δ_1,Δ_2) are given in (8)(9). Here we give the detailed definitions of the erased symbol sequences $X_{\mathsf{e}}^{\mathsf{I}}$ and $X_{\mathsf{e}}^{\mathsf{II}}$ in Phase III. First, the $X_{\mathsf{R}}[t]$, $t=1\ldots T$ in Phase I forms a codeword from a random Gaussian codebook to encode Σ_1 for B_1. In Phase III, from the delayed state information, the relay knows the erased indexes ts where state sequence with length T_1 of which $(S_1[t],S_2[t])=(0,1)$ in Phase I, $1\leq t\leq T$. For each $X_{\mathsf{R}}[t]$, we assign it to an unique symbol $X_{\mathsf{e}}^{\mathsf{I}}[t']=X_{\mathsf{R}}[t]$ in the erased symbol sequence. If $T_1>Tp(1-p)$, then we abandon the last $T_1-Tp(1-p)$ symbols. On the contrary, if $T_1<Tp(1-p)$, we set $X_{\mathsf{e}}^{\mathsf{I}}[t']=0, t'=T_1+1,\ldots Tp(1-p)$. Then the total length of erased symbol sequence $X_{\mathsf{e}}^{\mathsf{I}}[t']$ in Phase I is $Tp(1-p)$. B_1 also knows the mapping from t to t' as Π_{I}, where $t=\Pi_{\mathsf{I}}(t')$ with $t'=1\ldots\min\{T_1,Tp(1-p)\}, 1\leq t\leq T$. The mapping Π_{I} is also known at receiver B_1. In Phase II, $X_{2\mathsf{R}}[t]$, $t=T+1,\ldots,2T$ (independent of $X_{1\mathsf{R}}[t]$ in (18)) forms a codeword from a Gaussian codebook to encode Σ_2 for B_2. And as aforementioned, we can form length $Tp(1-p)$ erased symbol sequence in Phase II $X_{\mathsf{e}}^{\mathsf{II}}[\Pi_{\mathsf{II}}(t')]=X_{1\mathsf{R}}[t]+X_{2\mathsf{R}}[t]$ from the state sequence with length T_2 of which $(S_1[t],S_2[t])=(1,0)$, $T+1\leq t\leq 2T$. The corresponding mapping Π_{II} is known at receiver B_2.

In Phase III, the relay compresses the sum of erased symbol sequence $X_{\mathsf{e}}^{\mathsf{I}}[t'] + X_{\mathsf{e}}^{\mathsf{II}}[t'], t' = 1 \ldots Tp(1-p)$ with successive refinement, and send the corresponding quantization indexes i_{c} and i_{r} through the on/off downlink. At receiver $\mathsf{B}_i, i = 1, 2$, with successfully decoding the quantization index(es), we wish to obtain the following reconstruction

$$X_{\mathsf{e}}^{\mathsf{I}}[t'] + X_{\mathsf{e}}^{\mathsf{II}}[t'] + Z_{\mathsf{D}i}[t'] \tag{33}$$

where $Z_{\mathsf{D}i} \sim \mathcal{CN}(0, \mathsf{D}_i)$ and $\mathsf{D}_1 \leq \mathsf{D}_2$. To solve this two-receiver source-channel coding problem, we need to carefully select the lengths and rates of the source and channel codes. The length of source code is $Tp(1-p)$, and the rate selection of the source code comes as follows. To validate (33), we form the following test channels for successive refinement

$$V_1 = X_{\mathsf{sum}} + Z; \tag{34}$$

$$V_2 = V_1 + Z' = X_{\mathsf{sum}} + Z + Z', \tag{35}$$

with source X_{sum} having variance 2 and same distribution as $X_{\mathsf{e}}^{\mathsf{I}}[t'] + X_{\mathsf{e}}^{\mathsf{II}}[t']$; reconstruction errors $Z \sim \mathcal{CN}(0, \mathsf{D}_1)$ and $Z' \sim \mathcal{CN}(0, \mathsf{D}_2 - \mathsf{D}_1)$. Here X_{sum}, Z and Z' are independent. Note that due to the $\mod L$ operation in (18), the source X_{sum} is not Gaussian. We choose the rate for common index i_{c} from (35) as

$$\begin{aligned}\log\left(\frac{2 + (\mathsf{D}_2 - \mathsf{D}_1) + \mathsf{D}_1}{(\mathsf{D}_2 - \mathsf{D}_1) + \mathsf{D}_1}\right) &= \log\left(1 + \frac{2}{\mathsf{D}_2}\right) \\ \geq I(X_{\mathsf{sum}}; V_2),\end{aligned} \tag{36}$$

where the inequality comes from that Gaussian source is the hardest one to quantize [12]. The rate for refinement index i_{r} is chosen as

$$\log\left(\frac{1 + \frac{2}{\mathsf{D}_1}}{1 + \frac{2}{\mathsf{D}_2}}\right) \geq I(X_{\mathsf{sum}}; V_1) - I(X_{\mathsf{sum}}; V_2). \tag{37}$$

Note that the RHS is similar to the achievable rate in a Gaussian wiretap channel. Thus from [16], we know that for any source with variance 2, the Gaussian source $\mathcal{CN}(0, 2)$ maximizes the RHS. Note that $V_2 \to V_1 \to X_{\mathsf{sum}}$. By choosing V_1 as the reconstruction distribution at B_1 while V_2 as that at B_2, from (34)-(37), the MSE distortion at B_1 is D_1 while that at B_2 is D_2 respectively [15].

Now we must ensure that the receiver B_1 can correctly decode both i_{c} and i_{r} while receiver B_2 can correctly decode i_{c}. This task is done by carefully choosing the length and power allocation of the superposition channel coding, as well as D_1 and D_2 in (36)(37). First, since the channel has on/off probability p, it is crucial to choose the length of channel code longer than that of source code, which is $T(1-p)$ for the length $Tp(1-p)$ source code. Now indexes i_{c} and i_{r} are channel encoded using independent Gaussian codebooks with power $\mathsf{SNR}_{\mathsf{c}}$ and $\mathsf{SNR}_{\mathsf{r}}$ respectively, with power allocation $\mathsf{SNR}_{\mathsf{r}} = 1/\mathsf{SNR}_{\mathsf{2R}}, \mathsf{SNR}_{\mathsf{c}} = 1 - \mathsf{SNR}_{\mathsf{r}}$. For B_2 to correctly decode i_{c}, from (36) and the lengths of channel and source codes, we need

$$\begin{aligned}&Tp(1-p)\log\left(1 + \frac{2}{\mathsf{D}_2}\right) \\ \leq &T(1-p) \cdot p\, \mathcal{C}\left(\frac{\mathsf{SNR}_{\mathsf{2R}}\left(1 - \frac{1}{\mathsf{SNR}_{\mathsf{2R}}}\right)}{\mathsf{SNR}_{\mathsf{2R}}\frac{1}{\mathsf{SNR}_{\mathsf{2R}}} + 1}\right),\end{aligned}$$

which results in

$$\log\left(1 + \frac{2}{\mathsf{D}_2}\right) \leq \log\left(\frac{1 + \mathsf{SNR}_{\mathsf{2R}}}{2}\right). \tag{38}$$

Then we can choose

$$\mathsf{D}_2 = \frac{4}{\mathsf{SNR}_{\mathsf{2R}} - 1} \tag{39}$$

For receiver B_1, first note that since $\mathsf{SNR}_{\mathsf{1R}} \geq \mathsf{SNR}_{\mathsf{2R}}$, then B_1 can also successfully decode common index i_{c} by treating the codeword for i_{r} as noise. After subtracting the codewords corresponding to i_{c}, from (37) and the lengths of channel and source codes, we need

$$\log\left(1 + \frac{2}{\mathsf{D}_1}\right) - \log\left(1 + \frac{2}{\mathsf{D}_2}\right) \leq \mathcal{C}\left(\frac{\mathsf{SNR}_{\mathsf{1R}}}{\mathsf{SNR}_{\mathsf{2R}}}\right) \tag{40}$$

to correctly decode refinement index i_{r}. From (39), we must choose D_1 satisfying

$$\log\left(1 + \frac{2}{\mathsf{D}_1}\right) \leq \log\left(\frac{1 + \mathsf{SNR}_{\mathsf{2R}}}{2}\right) + \log\left(1 + \frac{\mathsf{SNR}_{\mathsf{1R}}}{\mathsf{SNR}_{\mathsf{2R}}}\right),$$

which is equivalent to

$$1 + \frac{2}{\mathsf{D}_1} \leq \frac{1}{2}\left(1 + \frac{\mathsf{SNR}_{\mathsf{1R}}}{\mathsf{SNR}_{\mathsf{2R}}} + \mathsf{SNR}_{\mathsf{2R}} + \mathsf{SNR}_{\mathsf{1R}}\right).$$

Because $\mathsf{SNR}_{\mathsf{1R}}/\mathsf{SNR}_{\mathsf{2R}} \geq 1$, the above inequality can be meet if

$$\mathsf{D}_1 = \frac{4}{\mathsf{SNR}_{\mathsf{1R}} + \mathsf{SNR}_{\mathsf{2R}}} \tag{41}$$

Now receivers $\mathsf{B}_i, i = 1, 2$ can obtain reconstructions (33) with D_1 in (41) and D_2 in (39) respectively, where D_1 in (39) is smaller than D_2 in (39). Note that the (noisy) erased sequence in Phase I $X_{\mathsf{e}}^{\mathsf{I}}[t']$ is known at B_2. As a side-information, B_2 can subtract $X_{\mathsf{e}}^{\mathsf{I}}[t']$ from reconstructions (33) and obtain $X_{\mathsf{e}}^{\mathsf{II}}[t'] + Z_{\mathsf{D}2}[t']$ (with additional channel noise). Now B_2 can combine the erased $X_{\mathsf{e}}^{\mathsf{II}}[t'] + Z_{\mathsf{D}2}[t']$ with the un-erased symbols received in Phase II to decode XOR Σ_2, with bounded gap to (16) as

$$R_2 \geq \frac{p(2-p)}{3-p}\mathcal{C}(\mathsf{SNR}_{\mathsf{2R}}) - \frac{p(1-p)}{3-p}\log(10) - \frac{p}{3-p}. \tag{42}$$

The details come as follows. First, we can upper-bound the MSE distortion D_2 in (39) as

$$\mathsf{D}_2 = \frac{4}{\mathsf{SNR}_{\mathsf{2R}} - 1} \leq \frac{4}{\frac{1}{2}\mathsf{SNR}_{\mathsf{2R}}} = \frac{8}{\mathsf{SNR}_{\mathsf{2R}}}. \tag{43}$$

The above inequality is due to that the SNR regime we considered $\mathsf{SNR}_{\mathsf{2R}} \geq 2$ is equivalent to $\mathsf{SNR}_{\mathsf{2R}} - 1 \geq \frac{1}{2}\mathsf{SNR}_{\mathsf{2R}}$. And from test channel (35), it is ensured that the quantization noise $Z_{\mathsf{D}2}[t']$ is independent of $X_{\mathsf{e}}^{\mathsf{II}}[t']$. Now we can from the following sequence to decode Σ_2 at receiver B_2

$$\tilde{Y}_{\mathsf{B}_2}[t] = \begin{cases} X_{\mathsf{R}}[\Pi_{\mathsf{II}}^{-1}(t')] + Z_{\mathsf{D}_2}[\Pi_{\mathsf{II}}^{-1}(t')] - \frac{Z_{\mathsf{B}_2}[t]}{h_{\mathsf{B}_2\mathsf{R}}} & t = \Pi_{\mathsf{II}}(t') \\ S_2[t] h_{\mathsf{B}_2\mathsf{R}} X_{\mathsf{R}}[t] + Z_{\mathsf{B}_2}[t] & \text{other } t, \end{cases}$$

where $X_{\mathsf{R}}[t] = X_{2\mathsf{R}}[t] + X_{1\mathsf{R}}[t]$ with $X_{2\mathsf{R}}[t]$ carrying message Σ_2 and interference $X_{1\mathsf{R}}[t]$ from (18), $T+1 \le t \le 2T$, $t' = 1 \dots \min\{T_1, T_2, Tp(1-p)\}$. Note that when $T \to \infty$, $T_1/T, T_2/T \to p(1-p)$ almost surely from our random state process. The power allocations of $X_{1\mathsf{R}}$ and $X_{2\mathsf{R}}$ are $1/\mathsf{SNR}_{2\mathsf{R}}$ and $1 - 1/\mathsf{SNR}_{2\mathsf{R}}$. When $T \to \infty$, from the independence of $X_{2\mathsf{R}}[t]$ and $X_{1\mathsf{R}}[t]$ and our power allocation, we have the following achievable rate for user B_2

$$R_2 \ge \frac{1}{(3-p)}\left(p(1-p)\mathcal{C}\left(\frac{1-\frac{1}{\mathsf{SNR}_{2\mathsf{R}}}}{\frac{8}{\mathsf{SNR}_{2\mathsf{R}}}+\frac{1}{\mathsf{SNR}_{2\mathsf{R}}}+\frac{1}{\mathsf{SNR}_{2\mathsf{R}}}}\right) + p\,\mathcal{C}\left(\frac{1-\frac{1}{\mathsf{SNR}_{2\mathsf{R}}}}{\frac{1}{\mathsf{SNR}_{2\mathsf{R}}}+\frac{1}{\mathsf{SNR}_{2\mathsf{R}}}}\right)\right) \quad (44)$$

where (44) comes from (43) and the fact that Gaussian interference is the worst interference under the same power constraint $1/\mathsf{SNR}_{2\mathsf{R}}$. Then (42) can be easily obtained from (44).

Now we show that for user B_1, bounded-gap rate $R_1 - \Delta_1$ with R_1 in (15) and Δ_1 in (8), is achievable. Following similar procedure as B_2 aforementioned, by combining the erased symbols with the un-erased symbols received in Phase I, the following rate is achievable to decode XOR Σ_1,

$$\frac{1}{(3-p)}\left(p(1-p)\mathcal{C}\left(\frac{1}{\frac{2}{\mathsf{SNR}_{1\mathsf{R}}}+\frac{1}{\mathsf{SNR}_{1\mathsf{R}}}}\right) + p\,\mathcal{C}(\mathsf{SNR}_{1\mathsf{R}})\right) \quad (45)$$

$$\ge \frac{p(2-p)}{3-p}\mathcal{C}(\mathsf{SNR}_{1\mathsf{R}}) - \frac{p(1-p)}{3-p}\log(3). \quad (46)$$

where the following inequality (from (41)) is used for obtaining (45)

$$\mathsf{D}_1 = \frac{4}{\mathsf{SNR}_{1\mathsf{R}} + \mathsf{SNR}_{2\mathsf{R}}} \le \frac{2}{\mathsf{SNR}_{1\mathsf{R}}},$$

which follows from the assumption $\mathsf{SNR}_{1\mathsf{R}} \ge \mathsf{SNR}_{2\mathsf{R}}$. Moreover, user B1 can decode additional messages by forming the following channel from the un-erased symbols in Phase II,

$$\begin{aligned}&\{w\,[h_{\mathsf{B}_1\mathsf{R}}\,(X_{1\mathsf{R}}[t] + X_{2\mathsf{R}}[t]) + Z_{\mathsf{B}_1}[t]] + d(t)\} \mod L\\ &= [C_{1\mathsf{R}}[t] + E_L(t)] \mod L, \quad (47)\end{aligned}$$

where $E_L(t) = (wh_{\mathsf{B}_1\mathsf{R}}-1)X_{1\mathsf{R}}[t] + wZ_{\mathsf{B}_1}[t]$. The channel (47) is a modulo-L channel with length Tp and power $L^2/12 = 1/\mathsf{SNR}_{2\mathsf{R}}$. By choosing w as the MMSE coefficient, rate

$$\frac{p}{3-p}\left(\log\left(\frac{\mathsf{SNR}_{1\mathsf{R}}}{\mathsf{SNR}_{2\mathsf{R}}}\right) - \log(2\pi e/12)\right) \quad (48)$$

is achievable [14]. To compare with (15), note that the RHS of (15) can be rewritten as

$$\begin{aligned}&\frac{1}{3-p}\left(p\left(\mathcal{C}\left(\mathsf{SNR}_{1\mathsf{R}}\right) - \mathcal{C}\left(\mathsf{SNR}_{2\mathsf{R}}\right)\right) + p(2-p)\mathcal{C}\left(\mathsf{SNR}_{1\mathsf{R}}\right)\right)\\ &\le \frac{1}{3-p}\left(p\log\left(\frac{\mathsf{SNR}_{1\mathsf{R}}}{\mathsf{SNR}_{2\mathsf{R}}}\right) + p(2-p)\mathcal{C}\left(\mathsf{SNR}_{1\mathsf{R}}\right)\right), \quad (49)\end{aligned}$$

where the second inequality is due to assumption $\mathsf{SNR}_{1\mathsf{R}} \ge \mathsf{SNR}_{2\mathsf{R}}$. By summing (48) and (46), our achievable rate for user B_1 has bounded gap to (15) as

$$\begin{aligned}R_1 \ge\, &p\left(\mathcal{C}(\mathsf{SNR}_{1\mathsf{R}}) - \mathcal{C}(\mathsf{SNR}_{2\mathsf{R}})\right) + \frac{p(2-p)}{3-p}\mathcal{C}(\mathsf{SNR}_{2\mathsf{R}})\\ &- \frac{p(1-p)}{3-p}\log(3) - \frac{p}{3-p}\log\left(\frac{2\pi e}{12}\right).\end{aligned}$$

Finally, the intersection of (7) and (5) is

$$(R_1, R_2) = \Big(p\left(\mathcal{C}(\mathsf{SNR}_{1\mathsf{R}}) - \mathcal{C}(\mathsf{SNR}_{2\mathsf{R}}\right), p\mathcal{C}(\mathsf{SNR}_{2\mathsf{R}}\Big). \quad (50)$$

For this point, we use Gaussian superposition coding to transmit messages Σ_1 and Σ_2 with corresponding power allocations $1/\mathsf{SNR}_{2\mathsf{R}}$ and $1 - 1/\mathsf{SNR}_{2\mathsf{R}}$. Then we have the following achievable rate pair (R_1, R_2)

$$\Big(p\left(\mathcal{C}(\mathsf{SNR}_{1\mathsf{R}}) - \mathcal{C}(\mathsf{SNR}_{2\mathsf{R}}\right), p\mathcal{C}(\mathsf{SNR}_{2\mathsf{R}}) - p\Big),$$

which has bounded gap to (50). And it concludes our proof.

C. Detailed proof of the outer-bound region in (11)

Here we prove that with delayed state information, the region $\mathscr{R}^{\mathrm{dl}}_{\mathrm{out}}(\mathrm{d})$ defined from (5)(6)(7) is a outer-bound region for the Gaussian model (1)(2). We first focus on (7), which results from first forming an equivalent degraded downlink and then outer-bounding carefully to avoid explicitly selecting the auxiliary random variable U. Since the noises at receivers in (1) are independent of the delayed state feedback, we can change Y_{B_1} and Y_{B_2} as

$$Y_{\mathsf{B}_1} = S_1 X_{\mathsf{R}} + \frac{Z_{\mathsf{B}_1}}{h_{\mathsf{B}_1\mathsf{R}}} \quad (51)$$

$$Y_{\mathsf{B}_2} = S_2 X_{\mathsf{R}} + \frac{Z_{\mathsf{B}_2}}{h_{\mathsf{B}_2\mathsf{R}}} = S_2 X_R + \frac{Z_{\mathsf{B}_1}}{h_{\mathsf{B}_1\mathsf{R}}} + Z' \quad (52)$$

where $Z' \sim \mathcal{CN}(0, \frac{1}{\mathsf{SNR}_{2\mathsf{R}}} - \frac{1}{\mathsf{SNR}_{1\mathsf{R}}})$. By giving Y_{B_2} in (52) to B_1, we have physically degraded channel [12] $X_{\mathsf{R}} \to \{Y_{\mathsf{B}_1}, Y_{\mathsf{B}_2}\} \to Y_{\mathsf{B}_2}$ with the following outer bounds

$$\begin{aligned}R_2 &\le I(U; Y_{\mathsf{B}_2}|S) \quad &(53)\\ &= I(X_{\mathsf{R}}, U; Y_{\mathsf{B}_2}|S_2) - I(X_{\mathsf{R}}; Y_{\mathsf{B}_2}|U, S_2)\\ &= I(X_{\mathsf{R}}; Y_{\mathsf{B}_2}|S_2) - pr \quad &(54)\\ &= p \cdot \mathcal{C}(\mathsf{SNR}_{2\mathsf{R}}) - pr, \quad &(55)\end{aligned}$$

where $S = \{S_1, S_2\}$ in (53), and (54) comes from conditional Markov Chain $U \to X_{\mathsf{R}} \to Y_{\mathsf{B}_2}$ given S_2, and r is defined as

$$r \triangleq I(X_{\mathsf{R}}; Y_{\mathsf{B}_2}|U, S_2 = 1); \quad (56)$$

also

$$\begin{aligned}R_1 \le\, &I(X_{\mathsf{R}}; Y_{\mathsf{B}_1}, Y_{\mathsf{B}_2}|S, U)\\ =\, &I(X_{\mathsf{R}}; Y_{\mathsf{B}_1}|S, U) + I(X_{\mathsf{R}}; Y_{\mathsf{B}_2}|Y_{\mathsf{B}_1}, S, U)\\ =\, &pI(X_{\mathsf{R}}; Y_{\mathsf{B}_1}|U, S_1 = 1) + p(1-p)I(X_{\mathsf{R}}; Y_{\mathsf{B}_2}|U, S_2 = 1)\\ &+ p^2 I(X_{\mathsf{R}}; Y_{\mathsf{B}_2}|Y_{\mathsf{B}_1}, U, S_1 = S_2 = 1). \quad (57)\end{aligned}$$

From (51) and (52),

$$I(X_{\mathsf{R}}; Y_{\mathsf{B}_2}|Y_{\mathsf{B}_1}, U, S_1 = S_2 = 1) = 0,$$

then

$$R_1 \le pI(X_{\mathsf{R}};Y_{\mathsf{B}_1}|U,S_1=1)+p(1-p)r \tag{58}$$

With r defined in (56),

$$\begin{aligned}&I(X_{\mathsf{R}};Y_{\mathsf{B}_1}|U,S_1=1)-r\\ =&h\left(X_{\mathsf{R}}+\frac{Z_{\mathsf{B}_1}}{h_{\mathsf{B}_1\mathsf{R}}}\Big|U\right)-h\left(\frac{Z_{\mathsf{B}_1}}{h_{\mathsf{B}_1\mathsf{R}}}\right)\\ &-h\left(X_{\mathsf{R}}+\frac{Z_{\mathsf{B}_2}}{h_{\mathsf{B}_2\mathsf{R}}}\Big|U\right)+h\left(\frac{Z_{\mathsf{B}_2}}{h_{\mathsf{B}_2\mathsf{R}}}\right) &(59)\\ \le&\mathcal{C}(\mathsf{SNR}_{1\mathsf{R}})-\mathcal{C}(\mathsf{SNR}_{2\mathsf{R}}). &(60)\end{aligned}$$

Note that the RHS of (59) is similar to the achievable rate in a Gaussian wiretap channel. From [16] and $|h_{\mathsf{B}_1\mathsf{R}}| \ge |h_{\mathsf{B}_2\mathsf{R}}|$, (60) is valid since the RHS of (59) is maximized when X_{R} is Gaussian conditioned on U. Substitute (60) into (58), we have

$$\begin{aligned}R_1 &\le p\left(r+\mathcal{C}(\mathsf{SNR}_{1\mathsf{R}})-\mathcal{C}(\mathsf{SNR}_{2\mathsf{R}})\right)+p(1-p)r\\ &=p(2-p)r+p\left(\mathcal{C}(\mathsf{SNR}_{1\mathsf{R}})-\mathcal{C}(\mathsf{SNR}_{2\mathsf{R}})\right)\end{aligned}$$

Together with (55), constraint (7) is obtained.

For (6), let us give Y_{B_1} in (51) to receiver B_2, we have a physically degraded channel $X_{\mathsf{R}} \to \{Y_{\mathsf{B}_1}, Y_{\mathsf{B}_2}\} \to Y_{\mathsf{B}_1}$, and having

$$R_1 \le p\,\mathcal{C}(\mathsf{SNR}_{1\mathsf{R}})-pr', \tag{61}$$

where $r'=I(X_{\mathsf{R}};Y_{\mathsf{B}_1}|U,S_1=1)$. This inequality follows from steps to reach (55). Also following steps to reach (57)

$$\begin{aligned}R_2 &\le pr'+p^2I(X_{\mathsf{R}};Y_{\mathsf{B}_2}|Y_{\mathsf{B}_1},U,S_1=1,S_2=1)\\ &\quad+p(1-p)(X_{\mathsf{R}};Y_{\mathsf{B}_2}|U,S_2=1,S_1=0)\\ &\le pr'+p^2\cdot 0+p(1-p)r' &(62)\\ &=pr'(2-p), &(63)\end{aligned}$$

where (62) comes from the data processing inequality,

$$I(X_{\mathsf{R}};Y_{\mathsf{B}_2}|U,S_2=1) \le I(X_{\mathsf{R}};Y_{\mathsf{B}_1}|U,S_1=1)=r', \tag{64}$$

since given U and $S_1=S_2=1$, we have the Markov chain $X_{\mathsf{R}} \to Y_{\mathsf{B}_1} \to Y_{\mathsf{B}_2}$ from (51) and (52). From (61) and (63), we have (6), and then $\mathscr{R}^{\mathrm{dl}}_{\mathrm{out}}(\mathrm{d})$ is a capacity outer bound region. Note that for the binary expansion model in Sec. V-A, by giving Y_{B_2} to B_1, we get

$$\tfrac{R_1}{p(2-p)}+\tfrac{R_2}{p} \le \tfrac{1}{(2-p)}(n_{1\mathsf{R}}-n_{2\mathsf{R}})+n_{2\mathsf{R}}. \tag{65}$$

from aforementioned degraded channel arguments. By reversing the role of B_1 and B_2, one get

$$\tfrac{R_1}{p}+\tfrac{R_2}{p(2-p)} \le n_{1\mathsf{R}}, \tag{66}$$

And the corner point (17) comes from the intersection of (65) and (66).

The outer-bound region $\mathscr{R}^{\mathrm{ul}}_{\mathrm{out}}(\mathrm{d})$ can be proved by allowing users A_1 and A_2 cooperate, which is akin to a two transmitter-antennas MISO channel with per antenna power constraint. This concludes our proofs for outer bounds of capacity region with delayed state information $\mathscr{C}(\mathrm{d},\mathrm{d})$ in (11).

D. Proof for the rest three regions in Theorem 3.1

Now we turn to the bounded-gap result for the capacity region with instantaneous state information for all terminal users and the relay $\mathscr{C}_{\mathrm{G}}(\mathrm{i},\mathrm{i})$. For (10), to achieve $\mathscr{R}^{\mathrm{ul}}_{\mathrm{G,in}}(\mathrm{i})$ for the uplink phase, the operations of users and the relay are similar to those for achieving $\mathscr{R}^{\mathrm{ul}}_{\mathrm{G,in}}(\mathrm{d})$ with delayed state information in Sec IV. The only difference is that one can perform on/off power allocation on (14) with instantaneous state information. In the downlink phase, the region $\mathscr{R}^{\mathrm{dl}}_{\mathrm{G,in}}(\mathrm{i})$ is fully achievable by Gaussian superposition coding with on/off power allocation. As for the outer-bound regions in (11), the cooperative outer bounds for the uplink with on/off power allocation result in $\mathscr{R}^{\mathrm{ul}}_{\mathrm{G,out}}(\mathrm{i})$. From the uplink-downlink duality (sum power constraint), $\mathscr{R}^{\mathrm{dl}}_{\mathrm{G,out}}(\mathrm{i})$ also a capacity outer-bound region.

Note that in our aforementioned proofs for $\mathscr{C}_{\mathrm{G}}(\mathrm{d},\mathrm{d})$ and $\mathscr{C}_{\mathrm{G}}(\mathrm{i},\mathrm{i})$, the capacity outer and inner bounds can be decomposed to those for uplink and downlink. Then these proofs also apply to proving $\mathscr{C}_{\mathrm{G}}(\mathrm{i},\mathrm{d})$ and $\mathscr{C}_{\mathrm{G}}(\mathrm{d},\mathrm{i})$, which establishes the bounded-gap results for all four (u,r) combinations in Theorem 3.1.

REFERENCES

[1] W. Nam, S.-Y. Chung, and Y. H. Lee, "Capacity of the Gaussian two-way relay channel to within 1/2 bit," *IEEE Transactions on Information Theory*, vol. 56, no. 11, pp. 5488–5494, Nov. 2010.

[2] A. Sezgin, A. S. Avestimehr, M. A. Khajehnejad, and B. Hassibi, "Divide-and-conquer: Approaching the capacity of the two-pair bidirectional Gaussian relay network," *IEEE Transactions on Information Theory*, vol. 58, no. 4, pp. 2434 – 2454, April 2012.

[3] A. Osseiran *et al.*, "Scenarios for 5G mobile and wireless communications: the vision of the METIS project," *IEEE Communications Magazine*, vol. 52, no. 5, pp. 26–35, May 2014.

[4] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Approximate capacity of the two-user MISO broadcast channel with delayed CSIT," in *Allerton Conference*, Illinois, USA, Oct. 2013, pp. 54–61.

[5] A. S. Avestimehr, S. N. Diggavi, and D. N. C. Tse, "Wireless network information flow: a deterministic approach," *IEEE Transactions on Information Theory*, vol. 57, no. 4, pp. 1872 – 1905, April 2011.

[6] B. Nazer and M. Gastpar, "Compute-and-forward: Harnessing interference through structured codes," *IEEE Transactions on Information Theory*, vol. 57, no. 10, pp. 6463–6486, Oct. 2011.

[7] I.-H. Wang, C. Suh, S. Diggavi, and P. Viswanath, "Bursty interference channel with feedback," in *IEEE Int. Symp. Inf. Theory (ISIT)*, 2013.

[8] A. Vahid, M. A. Maddah-Ali, and A. S. Avestimehr, "Capacity results for binary fading interference channels with delayed CSIT," *IEEE Transactions on Information Theory*, vol. 60, no. 10, pp. 6093–6130, Oct. 2014.

[9] S. Hua, C. Geng, and S. A. Jafar, "Topological interference management with alternating connectivity," in *IEEE Int. Symp. Inf. Theory (ISIT)*, 2013.

[10] P. Minero, M. Franceschetti, and D. Tse, "Random access: An information-theoretic perspective," *IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 909–930, Feb. 2012.

[11] L. Georgiadis and L. Tassiulas, "Broadcast erasure channel with feedback-capacity and algorithms," in *Proc. Workshop Network Coding, Theory, Appl.*, Lausanne, Switzerland, Jun. 2009, pp. 54–61.

[12] A. El-Gamal and Y.-H. Kim, *Network Information Theory*. Cambridge University Press, 2011.

[13] C.-P. Lee, S.-C. Lin, H.-J. Su, and H. V. Poor, "Multiuser lattice coding for the multiple-access relay channel," *IEEE Transactions on Wireless Communications*, vol. 13, no. 7, pp. 3539–3555, Jul. 2014.

[14] R. Zamir, S. Shamai, and U. Erez, "Nested linear/lattice codes for structured multiterminal binning," *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1250–1276, June 2002.

[15] S. Cheng and Z. Xiong, “Successive refinement for the Wyner-Ziv problem and layered code design,” *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 3269–3281, 2005.

[16] S. Leung-Yan-Cheong and M. E. Hellman, “The Gaussian wire-tap channel,” *IEEE Transactions on Information Theory*, vol. 24, no. 4, pp. 451–456, 1978.